## Linkages between streamflow, climate, and catchment characteristics: a global analysis

Beck, H.E.

2013

# Chapter 2

# Improving Curve Number based storm runoff estimates using soil moisture proxies[1]

**Abstract.** Advances in data dissemination and the availability of new remote sensing datasets present both opportunities and challenges for hydrologists in improving flood forecasting systems. The current study investigates the improvement in SCS Curve Number (CN)-based storm runoff estimates obtained after inclusion of various soil moisture proxies based on additional data on precipitation, baseflow, and soil moisture. A dataset (1980–2007) comprising 186 Australian catchments (ranging from 51 to 1979 km$^2$ in size) was used. In order to investigate the value of a particular proxy, the observed $S$ was compared to values obtained with different soil moisture proxies using linear regression. An antecedent precipitation index (API) based on gauged precipitation using a decay parameter proved most valuable in improving storm runoff estimates, stressing the importance of high quality precipitation data. An antecedent baseflow index (ABFI) also performed well. Proxies based on remote sensing (TRMM and AMSR-E) gave promising results, particularly when considering the expected arrival of higher accuracy data from upcoming satellites. The 5-day API performed poorly. The inclusion of soil moisture proxies resulted in mean modeled vs. observed correlation coefficients around 0.75 for almost all proxies. The greatest improvement in runoff estimates was observed in drier catchments with low Enhanced Vegetation Index (EVI) and topographical slope (all intercorrelated parameters). The present results suggest the usefulness of incorporating remotely sensed proxies for soil moisture and catchment wetness in flood forecasting systems.

[1]This chapter is an edited version of: Beck, H. E.; de Jeu, R. A. M.; Schellekens, J.; Van Dijk, A. I. J. M., and Bruijnzeel, L. A. Improving curve number based storm runoff estimates using soil moisture proxies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(4):250–259, 2009.

## 2.1 Introduction

Floods are among the most costly and frequent natural disasters in terms of people's suffering and economic losses. In fact, flooding is considered to be one of the prime catastrophic events threatening society in many countries (NRC, 1996). Under global environmental changes as a consequence of land use modifications and increased greenhouse gas emissions the flooding risk may increase (Betts et al., 2007). Flood forecasting systems provide a tool that allows decision makers (and in some cases the general public) to be proactive rather than reactive to flood hazard events. Recent developments in remote sensing and data dissemination present both opportunities and challenges for hydrologists in improving currently used flood forecasting systems since remote sensing offers a means to provide frequent global coverage of such critical hydrological data as precipitation and soil moisture (Entekhabi et al., 1999). The anticipated availability of high-accuracy data at finer spatial and temporal resolutions from ESA's Soil Moisture and Ocean Salinity (SMOS) mission in 2009 (Kerr et al., 2001), NASA's Soil Moisture Active/Passive (SMAP) mission in 2012 (NRC, 2007b), and NASA's Global Precipitation Measurement (GPM) mission in 2013 (Smith et al., 2006) presents another incentive to investigate the application of currently available satellite data.

Most nations have organized and operate special *in-situ* networks devoted to measuring river discharge and precipitation. These networks are increasingly made available to the public via internet. Examples are the National Climatic Data Center (NCDC) for precipitation in the US (see `http://www.ncdc.noaa.gov`) and the Queensland Natural Resources and Mines (NR&M) Data Drill for precipitation in Australia (see `http://www.longpaddock.qld.gov.au/silo/`). Also increasingly available are datasets incorporating multiple networks, such as the Global Precipitation Climatology Centre (GPCC) database (Schneider et al., 2008), the Global Precipitation Climatology Project (GPCP) (Adler et al.,

2003), the Global River Discharge Database (GRDD) (Vörösmarty et al., 1996), and the Global Soil Moisture Data Bank (GSMDB) (Robock et al., 2000). Whilst many of these datasets have their limitations—including the use of many different data-formats, low spatial and/or temporal resolution, or a high cost—improving their accessibility would present major opportunities for hydrologists involved in flood forecasting. The Atmospheric Data Access for the Geospatial User Community (ADAGUC) web portal of the Royal Meteorological Institute of The Netherlands (see `http://adaguc.knmi.nl`) has been created specifically to address these issues and to provide easy access to various datasets in a standardized and user-friendly format. It hosts a series of remote sensing products including AMSR-E soil moisture which is used in the current study.

At the base of many flood forecasting exercises lies the need for a robust model to convert input precipitation to storm runoff from a catchment. The Natural Resources Conservation Service (NRCS, formerly the Soil Conservation Service, SCS) Curve Number (CN) method (USDA, 1986) is widely used to estimate runoff from rainfall amounts. The attraction of the CN model lies in its simplicity (it requires the estimation of a single parameter only, the Curve Number) and in its numerous applications since the 1980s (Choi et al., 2002; Arnold and Fohrer, 2005; Kim and Lee, 2008). Because soil moisture is a key factor determining the partitioning of rainfall into runoff and infiltration (Aubert et al., 2003), soil moisture proxies are calculated to account for a catchment's wetness status prior to the rainfall event in the hope to improve stormflow prediction (USDA, 1986). There is some controversy as to the degree of improvement in storm runoff estimates that may be obtained when using additional (satellite-based) information on soil moisture. Some studies that used only a small number of catchments reported strong improvements in runoff estimates after the addition of remotely sensed soil moisture data (Jacobs et al., 2003; Brocca et al., 2008). Others found an improvement only for a subset of their catchments (Parajka et al., 2005b; Pauwels et al., 2002). This paper investigates the degree in improvement in storm runoff prediction obtained with soil moisture proxies derived from gauged precipitation, gauged stream discharge, AMSR-E soil moisture, and TRMM precipitation. Data from a large number (186) of Australian catchments is used, representing a range of precipitation, streamflow and soil moisture conditions. Through analysis of such a large number of catchments some understanding may be gained as to under what conditions the additional information on soil moisture will improve the runoff estimates.

## 2.2 Data and methods

### 2.2.1 Catchment selection

A large dataset (1980–2007) of 186 Australian catchments having daily streamflow and precipitation observations provided by CSIRO Land and Water was used in the present analysis. The catchments range in size from 51–1979 km$^2$ (mean size 447 km$^2$). Most catchments are located in the southeastern part of Australia (see Fig. 2.1a). This region is characterized by a temperate oceanic climate with a wet winter and low summer rainfall, whereas the central part of Australia is arid to semi-arid; the northern part has dry winters and relatively wet summers (see `http://www.bom.gov.au/climate/`). Most catchments are water-limited (80% of the catchments has ratio of long-term precipitation to potential evaporation $< 1$), are dominated by extensive agriculture or native vegetation (mean vegetation cover $35\% \pm 23$), and have for the main part low-relief (only 38% of the catchments has an average slope $> 10°$). Catchments where streamflow was subject to regulation or diversion were not included in the dataset.

### 2.2.2 Streamflow and precipitation data

Time series of daily precipitation were based on a gridded dataset (see `http://www.bom.gov.au/silo/`) with a spatial resolution of 0.05° based on interpolation over approximately 7200 meteorological stations (see Fig. 2.1b) (Jeffrey et al., 2001). Mean catchment precipitation was calculated from the area-weighted mean precipitation of the cells covered by the catchment.

Time series of streamflow were collected as part of previous studies (Guerschman et al., 2008; Peel et al., 2000). The baseflow component of the discharge is estimated using the Eckhardt recursive filter (Eckhardt, 2008), which low-pass filters the hydrograph according to:

$$Q_{\mathrm{bf}}(t) = \frac{(1-\mathrm{BFI}_{\max})\alpha Q_{\mathrm{bf}}(t-1) + (1-\alpha)\mathrm{BFI}_{\max} Q(t)}{1-\alpha \mathrm{BFI}_{\max}}, \tag{2.1}$$

subject to $Q_{\mathrm{bf}}(t) \leq Q(t)$, where $Q_{\mathrm{bf}}$ is the baseflow [m$^3$ s$^{-1}$], $Q$ is the total discharge [m$^3$ s$^{-1}$], $\mathrm{BFI}_{\max}$ is the maximum value of the baseflow index that can be modeled by the algorithm, $\alpha$ is the baseflow recession constant, and $t$ is time [days]. The $\alpha$ parameter is derived from recession analysis by plotting $Q(t_i)$ vs. $Q(t_i+1)$ for discharge values $Q(t_i)$ that are part of a recession period of at least five days, identified by:

$$Q(t_i-2) > Q(t_i-1) > Q(t_i) > Q(t_i+1) > Q(t_i+2). \tag{2.2}$$

Assuming that the aquifer is a linear reservoir, theoretically all points should follow a straight line through the origin with slope equal to the recession parameter $\alpha$. In practice, however, there is some scatter in the points. It

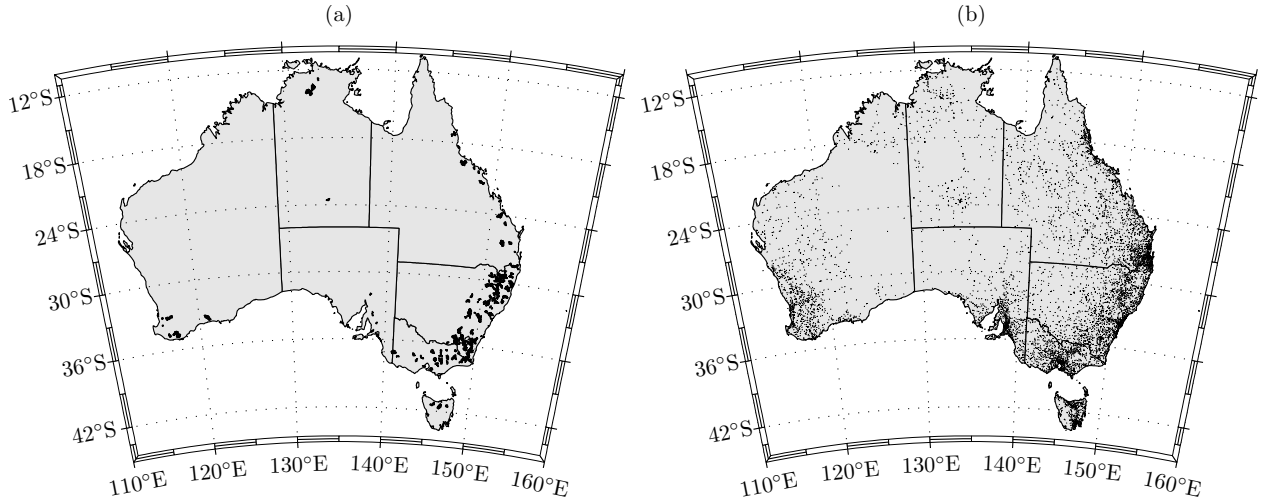


Figure 2.1: Map of Australia showing (a) the geographical distribution of the 186 catchments in the dataset, and (b) the locations of the precipitation gauges used on January 5, 2005. The actual gauges used vary from day to day depending on which are available, but will mostly be similar to this set.

is more probable that the points representing the slower recession represent the true value of $\alpha$, because (1) the discharge at $t_i$ could still contain direct runoff, and (2) there is probably more groundwater recharge at $t_i$. In reference (Eckhardt, 2008) the points representing the slowest recession (i.e., the upper limit of the $Q(t_i)$ vs. $Q(t_i + 1)$ plot) are on straight line with slope assumed to be $\alpha$. However, in the present study the points representing the slowest recession often approach one. To solve this problem common logarithms of both values are taken and $\alpha$ is determined by fitting a model of the form:

$$\log_{10}(Q(t_i)) = \alpha \log_{10}(Q(t_i + 1)) \tag{2.3}$$

through the points exceeding the $90^{\text{th}}$ percentile recession. Since most streams in southern Australia are ephemeral with porous aquifers, following recommendations in (Eckhardt, 2008) $\text{BFI}_{\text{max}}$ was set to 0.5 for all catchments. By using $\text{BFI}_{\text{max}} = 0.5$ Eq. 2.1 corresponds to the algorithm of Chapman and Maxwell (Chapman and Maxwell, 1996). Fig. 2.2 gives an example of baseflow separated using the described methodology for a single catchment (gauge code 410061) with $\alpha$ determined at 0.9823. The direct component of discharge ($Q_{\text{qf}}$) is calculated from $Q - Q_{\text{bf}}$.

Streamflow data represent totals based on the 24 hours prior to midnight local time, whereas the precipitation data are based on the 24 hours prior to 0900 local time. This 9 hour mismatch between precipitation and runoff days might result in a mismatch between $P$ events and their resulting $Q_{\text{qf}}$ in $9/24 = 38\%$ of the events (ignoring the delayed response of $Q_{\text{qf}}$). To reduce the effect of this mismatch the temporal resolution of $P$ and $Q_{\text{qf}}$ was resampled to 2 days (at 2 day intervals $t_i$ the mean of $t_i - 1$ and $t_i$ was calculated). This resulted in a decrease in the occurrence of mismatches between $P$ and corresponding $Q_{\text{qf}}$ to $9/48 = 19\%$ of the events. Catchments with less than 10 events with $P > 15$ mm after June 2002 (the launch of AMSR-E) were not included in the analysis.

### 2.2.3 AMSR-E soil moisture data

In this paper Advanced Microwave Scanning Radiometer Earth Observing System (AMSR-E)-based surface soil moisture is used to calculate a soil wetness index (SWI). AMSR-E Level-2A swaths from June 2002 (the launch of the AMSR-E mission) to present were used. Soil moisture values were derived using the Land Parameter Retrieval Model (LPRM) (Owe et al., 2008) from C-band (6.92 GHz) passive microwave brightness temperatures. AMSR-E soil moisture represents the soil moisture in the top 1–2 cm of the soil. Daily (24 hours prior to midnight) mean soil moisture was calculated from the average of all satellite observations in a swath with footprint centers inside the catchment boundaries during that day. The actual number of observations in a swath that were used thus depended on catchment size. On average, AMSR-E passes a location in Australia slightly more often than once per day. Both ascending (daytime, ±13:30 solar time) and descending (nighttime, ±1:30 solar time) passes of AMSR-E were used. Observations where the LPRM residual exceeded zero were excluded.

A validation study of the AMSR-E soil moisture product used in the present study was performed for Australia (Draper et al., 2007). A number of validation sites were in close proximity to the catchments used in the present study (although none were in densely vegetated regions). For these sites the normalized 5-day moving average of AMSR-E C-band ascending and descending soil moisture showed strong correlations with *in-situ* soil moisture time series at 0–7 cm depth (coefficient of determination $r^2 = 0.56$ to 0.87, RMSE = 0.016 to 0.066 $\text{m}^3\ \text{m}^{-3}$ volumetric soil moisture). In general there was good spatial agreement with antecedent precipitation. Radio frequency interference (RFI), a common problem in many areas in Europe and the US, is almost absent in Australia (Njoku et al., 2005).

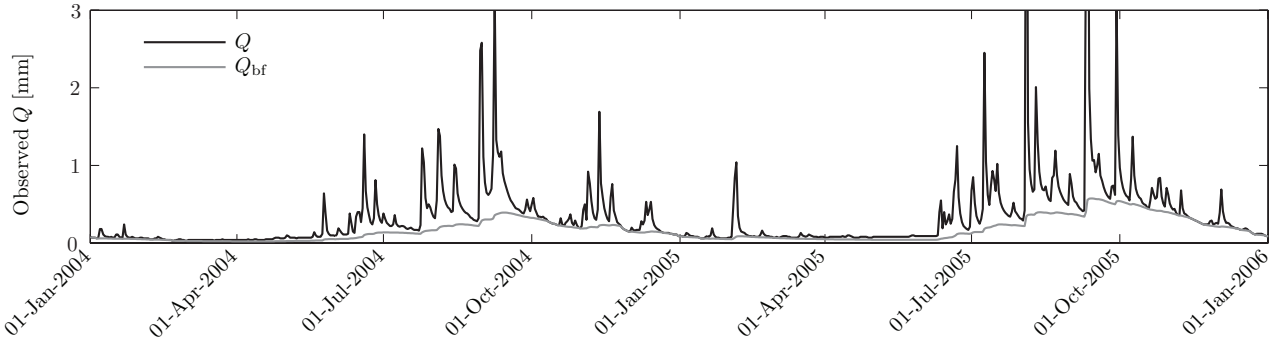


Figure 2.2: Baseflow separation results for catchment 410061 (gauge code). Baseflow was separated from total streamflow using the Eckhardt filter.

### 2.2.4 TRMM precipitation data

The current study uses the Tropical Rainfall Measuring Mission (TRMM) 3B42RT product, which is available at a lag of about 6 hours, and is a combination of TRMM real-time merged passive microwave (HQ, 3B40RT) and microwave-calibrated infrared (VAR, 3B41RT) (Huffman et al., 2003). For each pixel the HQ value is used if available, and otherwise the VAR value is used. The spatial resolution is 0.25°, the domain with useful data is 45°N–45°S, and the temporal resolution is 3 hours. Daily (24 hours prior to midnight local time) time series for each catchment were calculated from the area-weighted mean of all cells within the catchment. Compared to the TRMM 3B42 (research) product, the 3B42RT product does not include gauged precipitation data.

### 2.2.5 The Curve Number model

The Curve Number (CN) method is used to predict the amount of direct runoff resulting from large rainfall events in a particular area (USDA, 1986). The method is based on the following relationship:

$$\frac{F}{S} = \frac{Q_{\mathrm{qf}}}{P - I_a}, \tag{2.4}$$

where $F$ is the actual retention, $S$ the potential maximum retention, $Q_{\mathrm{qf}}$ the direct runoff, $P$ the precipitation, and $I_a$ the initial abstraction of rainfall by soil and vegetation (all in mm). The following continuity equation is introduced:

$$F = P - I_a - Q_{\mathrm{qf}} \tag{2.5}$$

Combining Eqs. 2.4 and 2.5 eliminates $F$, yielding:

$$Q_{\mathrm{qf}} = \frac{(P - I_a)^2}{P - I_a + S} \quad \text{for} \quad P > I_a. \tag{2.6}$$

$I_a$ is related to $S$ according to:

$$I_a = \lambda S, \tag{2.7}$$

where $\lambda$ is the initial abstraction coefficient.

The standard value for $\lambda$ is 0.2 (Ponce and Hawkins, 1996). However, several more recent studies have shown that the assumption of 0.2 is unusually high, and that values between 0.01 and 0.05 are more realistic. A study using rainfall and runoff data from 307 US catchments or plots found that a value of $\lambda$ of 0.05 would fit the data much better (Woodward et al., 2003). In an experimental watershed in the Three Gorges Area of China it was found that $\lambda$ varied from 0.010 to 0.154 with a median of 0.048 (Shi et al., 2009). In an experimental catchment in Attica, Greece, the average value of $\lambda$ was equal to 0.014, and in a subcatchment equal to 0.037 (Baltas et al., 2007). In yet another study using data from 237 US catchments optimal model performance was found with a value of $\lambda$ in the order of 0.01 (Mishra et al., 2004). In the present study a fixed value of 0.05 was used for all catchments.

To directly calculate $S$ for a given rainfall and runoff amount the following equation may be used (which is a combination of Eqs. 2.6 and 2.7, rewritten to solve for $S$):

$$\begin{aligned} S = 0.5\lambda^{-2}[&Q_{\mathrm{qf}} - \lambda Q_{\mathrm{qf}} + 2\lambda P \\ &- \sqrt{Q_{\mathrm{qf}}^2 - 2\lambda Q_{\mathrm{qf}}^2 + \lambda^2 Q_{\mathrm{qf}}^2 + 4\lambda P Q_{\mathrm{qf}}}]. \end{aligned} \tag{2.8}$$

Fig. 2.4b gives an example of $S$ values calculated using Eq. 2.8 ($\lambda = 0.05$) from observed rainfall and direct runoff data shown in Fig. 2.4a.

Because runoff usually varies widely for the same rainfall amount (mainly due to antecedent soil moisture conditions, but also due to spatial variability of rainfall, variability in intensity, measurement errors, and other factors) the observed $S$ varies between storms. The $S$ also appears to be biased towards high values at low storm depths (Hawkins, 1993) because there is not enough runoff. This effect is more evident at higher values of $\lambda$ because there has to be more rainfall for runoff to occur. To correct this bias a model is fitted with the following equation (from (Hawkins, 1993), rewritten for $S$):

$$S_{\mathrm{fit}} = S_{\mathrm{inf}} - S_{\mathrm{inf}} \exp(-vP), \tag{2.9}$$

where $S_{\mathrm{fit}}$ is the fitted model and $P$ again the amount of precipitation [mm]. $S_{\mathrm{inf}}$ (the asymptotic $S$ [mm]) and $v$ are fitting parameters. The $S$ values with the bias

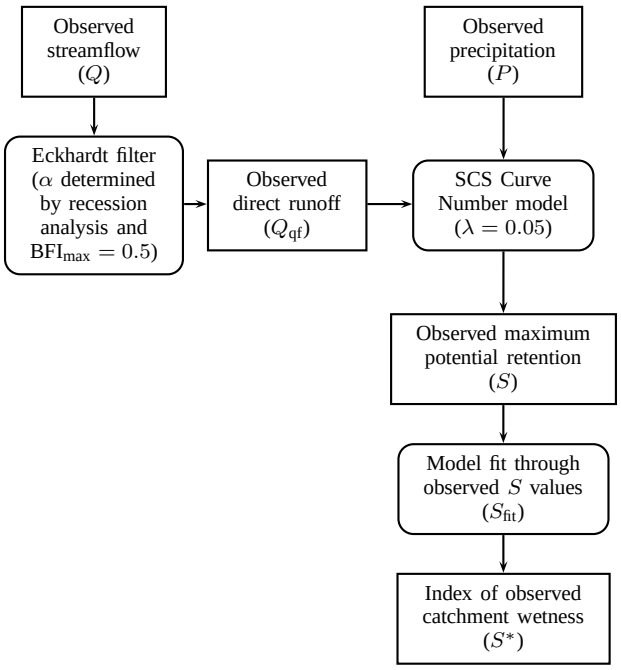


Figure 2.3: Flow-chart showing the derivation of $S^*$ for a single catchment.

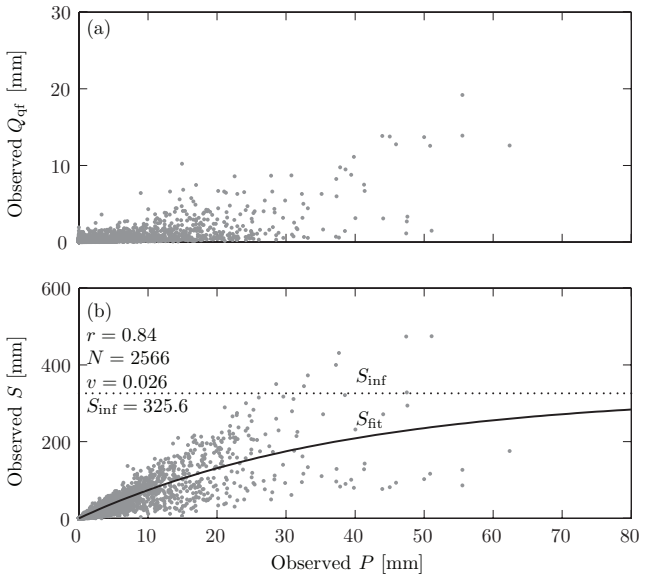


Figure 2.4: Plots for a single catchment (gauge code 410061) of (a) observed precipitation vs. observed direct runoff, and (b) observed precipitation vs. observed $S$ as calculated from observed precipitation and direct runoff using Eq. 2.8. Also shown is the fitted model ($S_{fit}$) and the asymptote of the fit ($S_{inf}$). The correlation coefficient $r$ of the fitted model and the number of observations $N$ are also listed. Data from 1980–2007 have been used.

removed are calculated as follows:

$$S^* = S - S_{\text{fit}}, \tag{2.10}$$

where $S^*$ may be interpreted as an index of mean catchment wetness or saturated area within a catchment. Ignoring other factors, a low $S^*$ value should indicate a wet catchment with a low infiltration rate and a reduced potential for storage, whereas a high $S^*$ value should indicate a drier catchment where a greater portion of the rainfall infiltrates. Fig. 2.3 summarizes the steps taken in calculating $S^*$ values for a catchment.

### 2.2.6 Soil moisture proxies

Soil moisture proxies are used to estimate the catchment wetness prior to a rainfall event and adjust Curve Number estimates of runoff accordingly. This section presents the five soil moisture proxies used in this paper (SWI, $\text{API}_k$, $\text{API}_{k\ \text{TRMM}}$, $\text{API}_{5\,\text{day}}$, and ABFI). The soil moisture proxies are calculated on a daily basis from June 2002 (the launch of the AMSR-E mission) until 2007 (marking the end of the precipitation and discharge time series).

#### Soil Wetness Index (SWI)

An exponential moving average filter (Wagner et al., 1999) is used to relate AMSR-E based surface soil moisture to the part of the soil that influences runoff production. The approach is based on the idea that surface soil moisture derived from AMSR-E data fluctuates at a higher frequency (since it pertains to the top few cm only) than the total runoff-generating part of the soil. The Soil Water Index (SWI [%]) is calculated according to:

$$\text{SWI}(t) = \frac{\sum_i \theta_{\text{AMSR-E}}(t_i) \exp\left(-\frac{t-t_i}{T}\right)}{\sum_i \exp\left(-\frac{t-t_i}{T}\right)}, \tag{2.11}$$

where $\theta_{\text{AMSR-E}}$ is the AMSR-E-based soil moisture content [% volumetric content] and $T$ is a time lag constant [days]. In the present study $T$ was optimized for each catchment seperately to give the highest linear correlation coefficient $r$ between antecedent SWI and $S^*$.

#### Gauge Antecedent Precipitation Index ($\text{API}_k$)

An antecedent precipitation index (API) is often used because it uses readily available precipitation data (Mishra et al., 2004). $\text{API}_k$ [mm] assumes a temporal decay constant $k$ to account for soil moisture losses. The following recursive model was used (Linsley et al., 1949):

$$\text{API}_k(t) = k\,\text{API}_k(t-1) + P(t), \tag{2.12}$$

where $P$ is the gauged precipitation [mm], and $k$ is a decay parameter ($< 1$) that is inversely related to drainage and evaporation losses and positively related to penetration depth [-]. In literature $k$ is reported to vary between 0.80 and 0.98 (Brocca et al., 2008). In the present study $k$ was optimized separately for each catchment based on the highest linear correlation coefficient $r$ between antecedent $\text{API}_k$ and $S^*$.

**TRMM Antecedent Precipitation Index ($\mathbf{API}_{k\ \mathrm{TRMM}}$)**

To investigate the performance of a soil moisture proxy based on near real-time satellite precipitation an API is calculated using Eq. 2.12 with TRMM precipitation estimates for $P$ and $k$ equal to the value optimized using gauged precipitation data.

**5-day Antecedent Precipitation Index ($\mathbf{API}_{5\,\mathrm{day}}$)**

Traditionally the amount of rainfall received in the five days preceding a storm event of interest is used (Chow et al., 1988). This 5-day API is calculated according to:

$$\mathrm{API}_{5\ \mathrm{day}}(t) = \sum_{i=0}^{4} P(t_i), \tag{2.13}$$

where $i$ denotes the $i$th day before day $t$.

**Antecedent Baseflow Index (ABFI)**

In literature an exponential relation between baseflow and groundwater storage has been proposed (Sivapalan et al., 1987). ABFI [-] is therefore calculated from the natural logarithm of the baseflow:

$$\mathrm{ABFI}(t) = \log_e\left(Q_{\mathrm{bf}}(t)\right), \tag{2.14}$$

where $Q_{\mathrm{bf}}$ is the baseflow [m$^3$ s$^{-1}$] and $t$ is time [days]. The advantage to this method is that it eliminates the need to select additional parameters. However, separation of baseflow from total flow is required. See section 2.2.2 for a description of the separation approach used here.

### 2.2.7 Approach

For each catchment, observed values of $S^*$ (derived from observed rainfall and direct runoff) for events with $P > 15$ mm were compared to antecedent values of the respective soil moisture proxies. $S^*$ values for day $t_i$ were compared to values of $\mathrm{API}_{k\ \mathrm{TRMM}}$, ABFI, and SWI on day $t_i - 2$ (due to the 2-day resampling of precipitation and direct runoff) and compared to values of $\mathrm{API}_k$ and $\mathrm{API}_{5\,\mathrm{day}}$ on day $t_i - 3$ (due to the 2-day resampling and the 9 hour mismatch between measurements of rainfall and streamflow). For each antecedent soil moisture proxy (ASMP) a linear regression is performed of the form:

$$S^* = a\mathrm{ASMP} + b, \tag{2.15}$$

where $a$ and $b$ are fit parameters. The quality of the correlation was taken to indicate the value of the soil moisture proxy for improving runoff estimates. An independent evaluation of the model was also performed using the same catchments but setting $T$ equal to 5 and $k$ equal to 0.971 for all catchments.

The improvement in model performance after including a soil moisture proxy in the CN model was assessed as follows. First, the correlation coefficient of modeled vs. observed $Q_{\mathrm{qf}}$ was calculated for the CN model without the inclusion of the proxy by using Eq. 2.6 with the $S$ values equal to $S_{\mathrm{fit}}$. Secondly, to quantify the improvement, the correlation coefficient of modeled vs. observed $Q_{\mathrm{qf}}$ was calculated again but now for the CN model with the proxy included by using Eq. 2.6 with $S$ calculated using Eqs. 2.9, 2.10, and 2.15.

## 2.3 Results and discussion

In the current study for 186 Australian catchments (51–1979 km$^2$) the performance of five soil moisture proxies in improving runoff estimates is investigated. Table 2.1 shows for each soil moisture proxy (1) the number of catchments having a significant correlation (probability $p < 0.01$) between antecedent values of the soil moisture proxy and $S^*$ (catchment wetness according to observed precipitation and runoff), (2) the mean correlation coefficient $r$, and (3) the obtained improvement in runoff estimates.

The $\mathrm{API}_k$ soil moisture proxy (with $k$ optimized) performed best, with a significant correlation obtained for 77% of the catchments. The ABFI and SWI (with $T$ optimized) proxies performed reasonably well, with significant correlations in 47% and 35% of the catchments, respectively. The $\mathrm{API}_{5\,\mathrm{day}}$ and $\mathrm{API}_{k\ \mathrm{TRMM}}$ proxies exhibited poorer performance with significant correlations in 22% and 23% of the catchments, respectively. The poor performance of $\mathrm{API}_{k\ \mathrm{TRMM}}$ may be due to the small size of the catchments, which on average cover less than a single TRMM cell. Better results can be expected by using the TRMM 3B42 research product, which includes gauge data, but these are not available at near real-time for operational systems. The good performance of $\mathrm{API}_k$ stresses the importance of high-quality precipitation data. All soil moisture proxies resulted in large improvements in modeled vs. observed runoff after including the proxy in the model, with correlation coefficients around 0.75 obtained for almost all proxies. Table 2.1 also lists results for $\mathrm{API}_k$ ($k = 0.971$), $\mathrm{API}_{k\ \mathrm{TRMM}}$ ($k = 0.971$), and SWI ($T = 5$ days), with slightly poorer results compared to the proxies with optimized parameters. Figs. 2.6, 2.7, and 2.8 show detailed results for a single catchment (gauge code 410061). In this catchment SWI performed particularly well, demonstrating the high potential of SWI as an index of overall catchment wetness status. Our results are similar to those of Parajka et al. (2005b) and Pauwels et al. (2002), who found inclusion of remotely sensed soil moisture improved the model in only a subset of the catchments, but also to Brocca et al. (2008) and Jacobs et al. (2003) who investigated a small number of catchments and found large improvements after the addition of remotely sensed soil moisture. The performance of the various proxies is not interchangeable (when for example ABFI performs well it is likely that $\mathrm{API}_k$ also performs well).

Table 2.1: Summary of the results showing the number of catchments with a significant correlation (probability $p < 0.01$) between the soil moisture proxy and $S^*$ (1), the mean correlation coefficient $r$ between the proxy and $S^*$ (3), and the improvement in runoff estimates after including the proxy into the model (4). The standard deviation of $r$ is also given.

| | Number of catchments | | Mean $Q_{qf}$ observed vs. modeled $r$ | |
|---|---|---|---|---|
| Soil moisture proxy | with $p < 0.01$ | Mean $r$ | Standard SCS-CN | SCS-CN with proxy |
| $API_{5\,day}$ | 41 (22%) | −0.62 (±0.12) | 0.68 (±0.26) | 0.80 (±0.17) |
| $API_k$ ($k$ optimized for each catchment) | 143 (77%) | −0.70 (±0.11) | 0.57 (±0.29) | 0.75 (±0.18) |
| $API_k$ ($k = 0.971$) | 110 (59%) | −0.68 (±0.11) | 0.57 (±0.28) | 0.73 (±0.17) |
| $API_{k\ TRMM}$ ($k$ equal to $API_k$ optimized $k$) | 42 (23%) | −0.62 (±0.11) | 0.69 (±0.23) | 0.79 (±0.17) |
| $API_{k\ TRMM}$ ($k = 0.971$) | 37 (20%) | −0.60 (±0.11) | 0.70 (±0.20) | 0.80 (±0.15) |
| ABFI | 88 (47%) | −0.69 (±0.12) | 0.55 (±0.30) | 0.76 (±0.16) |
| SWI ($T$ optimized for each catchment) | 65 (35%) | −0.71 (±0.12) | 0.47 (±0.30) | 0.72 (±0.16) |
| SWI ($T = 5$ days) | 47 (25%) | −0.68 (±0.12) | 0.43 (±0.30) | 0.67 (±0.16) |

The SWI time lag constant $T$ and the $API_k$ decay parameter $k$ were optimized for each catchment individually. Fig. 2.5 shows the optimization results for SWI and $API_k$. Only the catchments with a significant correlation between $S^*$ and antecedent values of the proxy are shown. The value of $T$ resulting in the best correlations was found to be 5 days, implying that 63% of the SWI values are based on data less than 5 days old. Higher values of $T$ were found in some other studies (Brocca et al., 2008; Wagner et al., 1999). The optimum value of $k$ was 0.971, implying that 63% of the $API_k$ values are based on rainfall less than 41 days old. A likely explanation for the difference in lag between SWI and $API_k$ is that remotely sensed soil moisture already contains information on past events.

An attempt was made to relate the correlation coefficient $r$ of the relationship between antecedent values of a soil moisture proxy and $S^*$ with catchment attributes. We examined catchment size, mean topographical slope, mean saturated conductivity, dominant texture class, mean plant available water content, mean clay content, mean precipitation, mean discharge, mean potential evaporation (ET), mean actual ET as derived from MODIS (Guerschman et al., 2008), mean Enhanced Vegetation Index (EVI) calculated from catchment average MOD34B4 reflectances, and fractions of woody vegetation (from NFI, 1997), non-agricultural land, grazing land, horticulture, and broad acre cropping. Correlations were found in turn with catchment mean precipitation, discharge, ET, topographical slope, and EVI (see Figs. 2.9, 2.10, and 2.11). For most soil moisture proxies and catchment parameters the correlations are significant at the 0.01 level. A study in Austria (Parajka et al., 2005b) found a similar relationship between the standard deviation of the catchment elevation (equivalent to topographical slope) and the correlation coefficient of the relationship between modeled and remotely sensed soil moisture. Due to intercorrelation between the respective parameters it is difficult to identify the primary mechanism behind these correlations. However, there are two possible explanations. Wet catchments (high mean precipitation, discharge, and ET) tend to have more frequent and more intense rainfall events that can be expected to exhibit a more straightforward relationship with runoff amount, leaving little room for improvement and thus resulting in lower correlation coefficients between antecedent values of the proxies and $S^*$. Fig. 2.12 demonstrates this by showing better model performance at higher ET in (a), and thus less improvement after the addition of a proxy into the model in (b). Poor model performance in drier catchments is also found in other studies (Parajka et al., 2005b; Lidén and Harlin, 2000; Gan and Biftu, 1996; Chiew et al., 1993). Another possibility is that the soil moisture in catchments with high topographical slope and high EVI is generally more spatially heterogeneous, thereby complicating the estimation of catchment-averaged soil moisture. This leads to a soil moisture proxy that is less reliable and thus to poorer performance.

Our results show that large improvements in Curve Number-based estimates of runoff are possible with the addition of soil moisture proxies for most catchments. The improvement is highest in drier catchments with low EVI and topographical slope. Although the remotely sensed proxies currently show lower performance, this could well improve in the near future with the availability of high-accuracy data at finer spatial and temporal resolutions from ESA's Soil Moisture and Ocean Salinity (SMOS) mission in 2009 (Kerr et al., 2001), NASA's Soil Moisture Active/Passive (SMAP) mission in 2012 (NRC, 2007b), and NASA's Global Precipitation Measurement (GPM) mission in 2013 (Smith et al., 2006).

## 2.4 Conclusions

The following conclusions can be drawn based on the analysis performed in the current study:

1. In 186 Australian catchments (51–1979 km$^2$) large improvements in the predictive capability of the

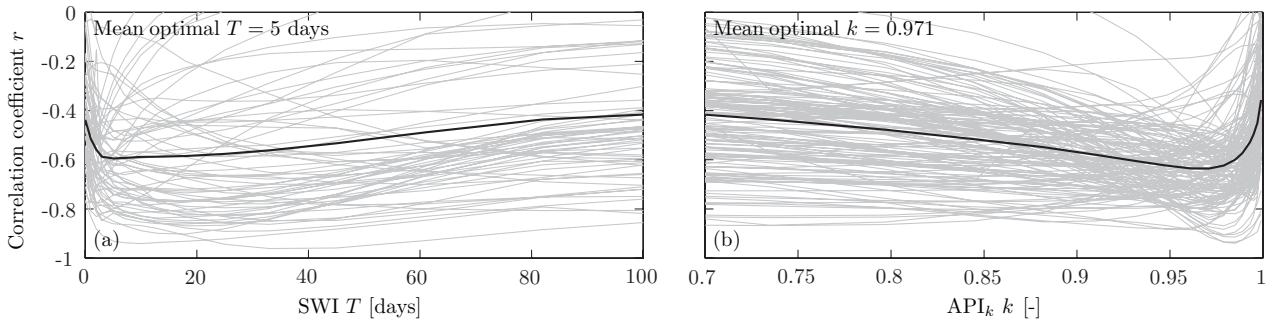


Figure 2.5: The correlation coefficient $r$ between $S^*$ and antecedent SWI for different values of $T$ is shown in (a). The correlation coefficient $r$ between $S^*$ and antecedent API$_k$ for different values of $k$ is shown in (b). Each gray line represents a catchment. The black line represents the mean over all catchments. Only catchments with a significant correlation have been used.

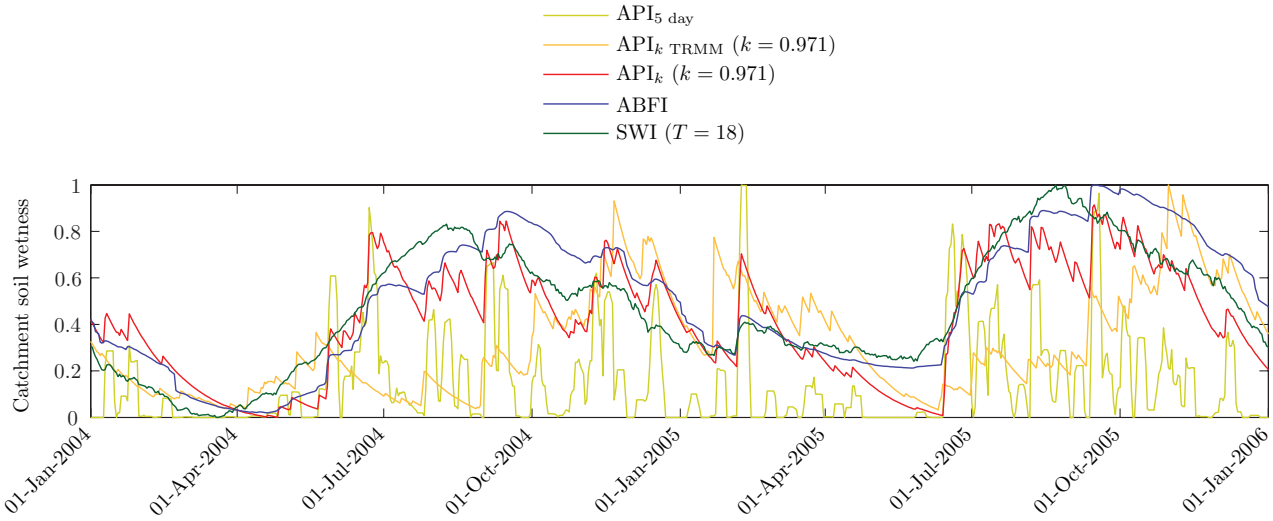


Figure 2.8: Results for catchment 410061 (gauge code) showing the temporal dynamics of the normalized soil moisture proxies.

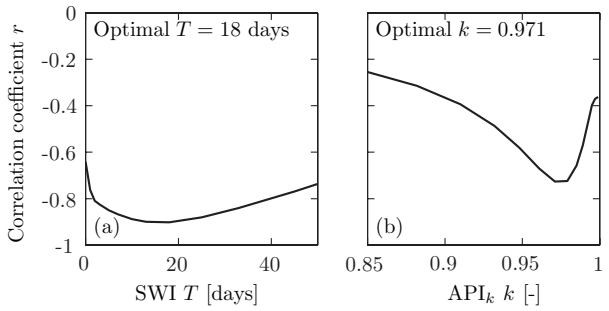


Figure 2.6: Results for catchment 410061 (gauge code) showing the optimization of (a) SWI $T$ and (b) API$_k$ $k$.

Curve Number model are possible by the use of soil moisture proxies derived from additional data on precipitation, soil moisture, and baseflow. Almost all soil moisture proxies were able to increase the mean modeled vs. observed runoff correlation to around 0.75.

2. API$_k$ performed best, stressing the importance of accurate precipitation data.
3. An optimal value of 5 days was found for the time lag constant $T$ used in the calculation of SWI, and an optimal value of 0.971 was found for the decay parameter $k$ used in the calculation of API$_k$.
4. In dry catchments with low EVI and topographical slope the addition of soil moisture proxies results in large improvements in runoff predictions. Due to covariances between the parameters distinct conclusions about the mechanism behind the correlations are speculative.
5. The soil moisture proxies derived from current satellites (SWI and API$_{k\ TRMM}$) show promising results.

## Acknowledgements

The authors wish to thank Juan Pablo Guerschman for assistance in developing the dataset used in this study.

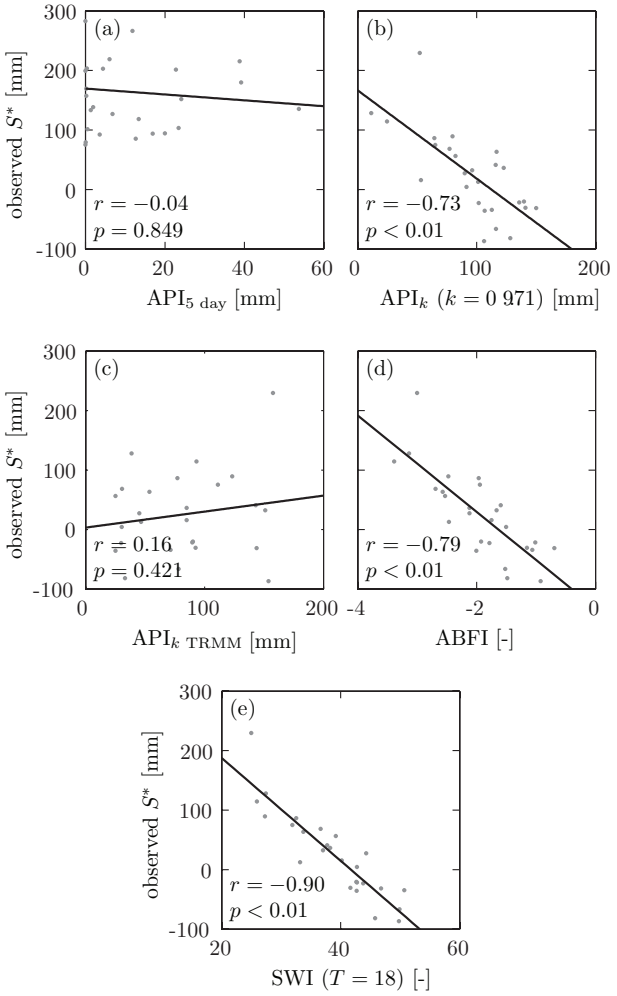


Figure 2.7: Results for catchment 410061 (gauge code) showing the regression between antecedent values of the respective soil moisture proxies and $S^*$.

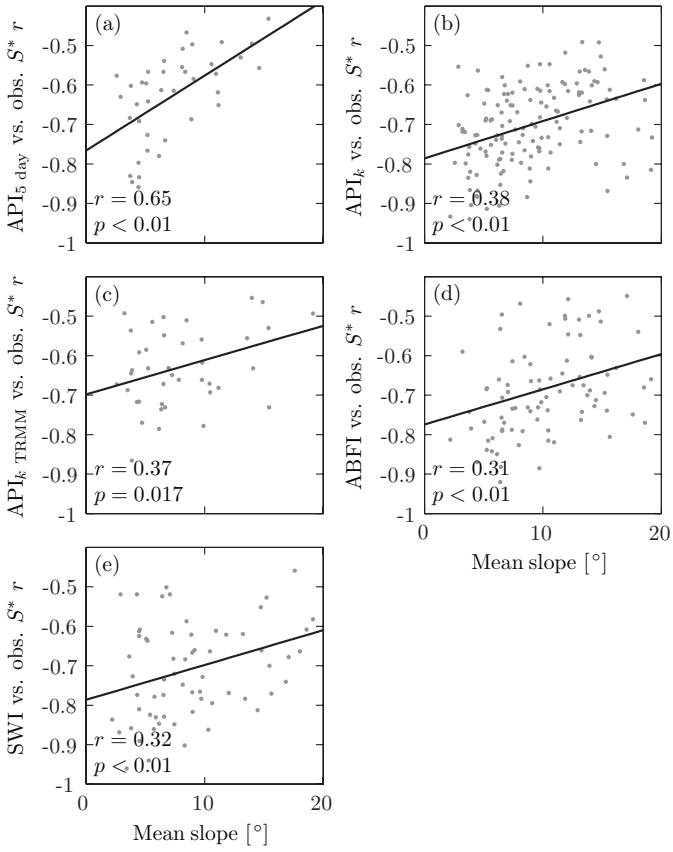


Figure 2.9: The performance of each soil moisture proxy plotted against catchment mean slope. Only the catchments with a significant correlation between the soil moisture proxy and $S^*$ are shown.

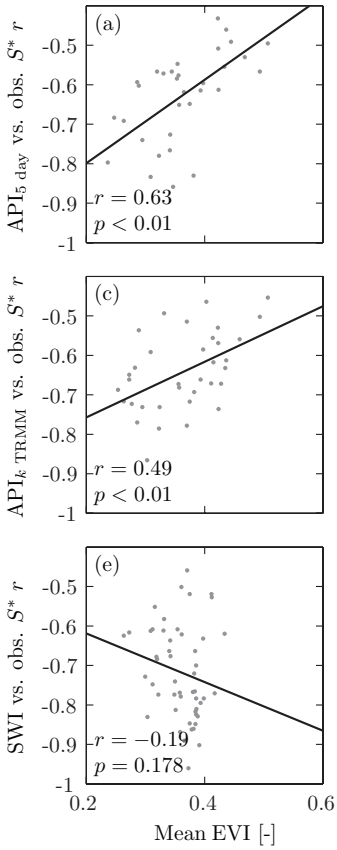


Figure 2.10: The performance of each soil moisture proxy plotted against catchment mean Enhanced Vegetation Index (EVI). Only the catchments with a significant correlation between the soil moisture proxy and $S^*$ are shown.

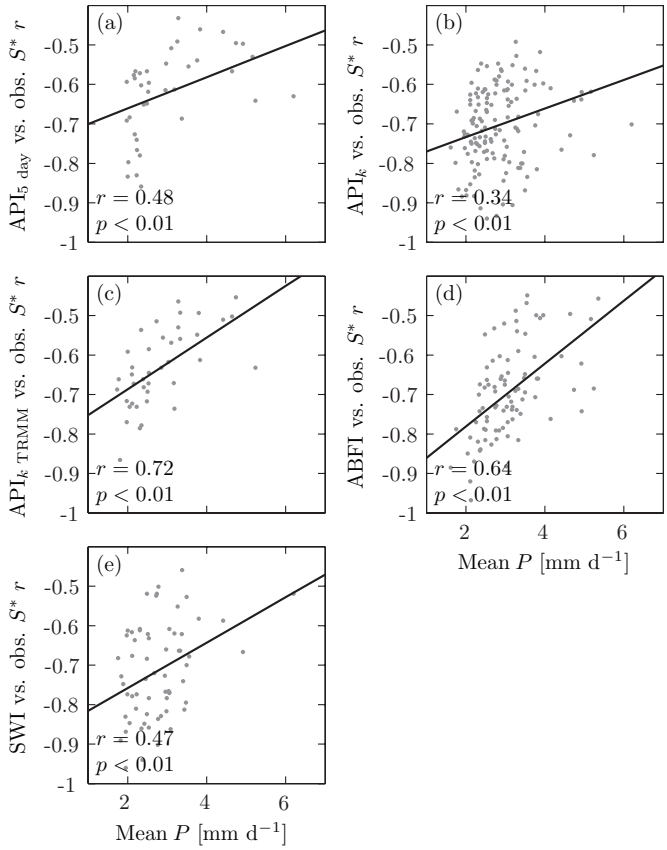


Figure 2.11: The performance of each soil moisture proxy plotted against catchment mean precipitation. Only the catchments with a significant correlation between the soil moisture proxy and $S^*$ are shown.

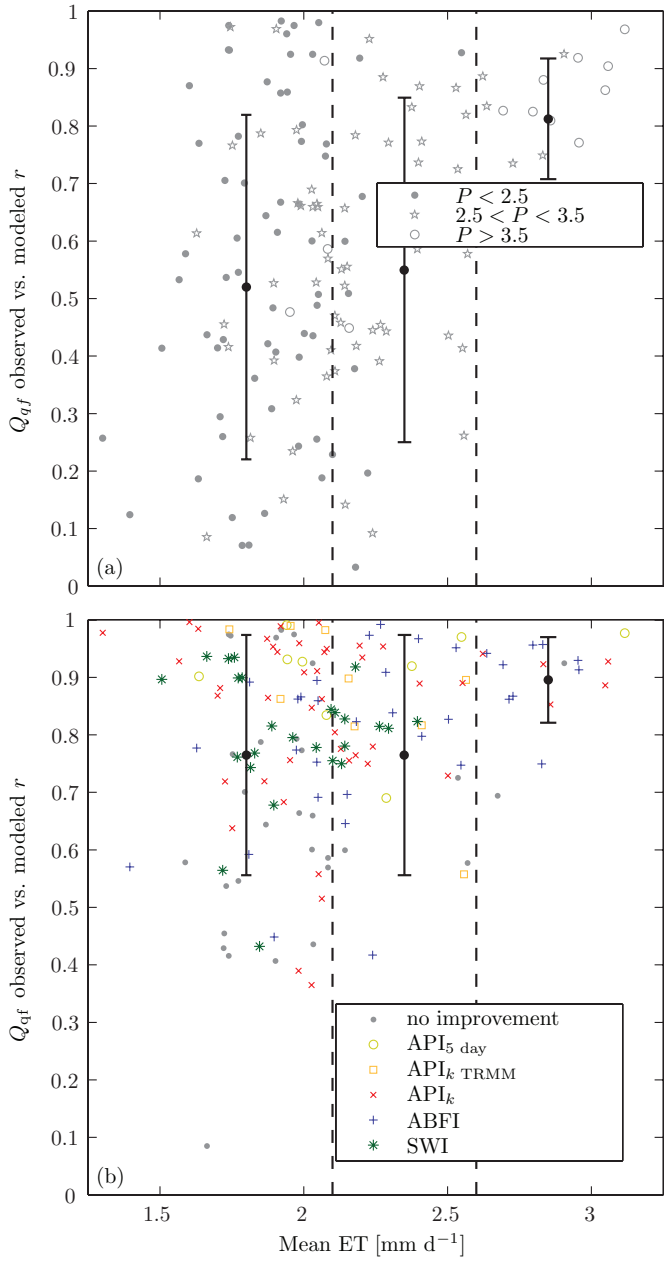


Figure 2.12: (a) Catchment mean evaporation (ET) plotted against the correlation coefficient $r$ of the relationship linking observed and modeled runoff. To show the relationship between ET and $P$ different symbols were used depending on the mean catchment $P$ [mm d$^{-1}$]. (b) Correlation coefficient $r$ of the relationship linking observed and modeled runoff after the inclusion of the soil moisture proxy that worked best. The error bars show the mean $r$ and standard deviation of the catchments between the dashed lines.